

データベース第8回

第8章 正規化理論 —高次の正規化—

高次の正規化: 更新時異状解消のため

- リレーションスキーマ中の属性に関する性質
 - 候補キー
 - 多値従属性
 - 関数従属性, 完全関数従属 など
- 正規形の定義
 - 上の性質を利用
 - 正規形を満たさないリレーションスキーマがあれば, 満たすように情報無損失に分解する=更新時異状解消

2

第2正規形

リレーションスキーマRが第2正規形であるとは次の2つの条件を満たすときをいう

1. Rは第1正規形である
2. Rの全ての非キー属性はRの各候補キーに**完全関数従属**している

注: 非キー属性とは, いかなる候補キーにも属していない属性

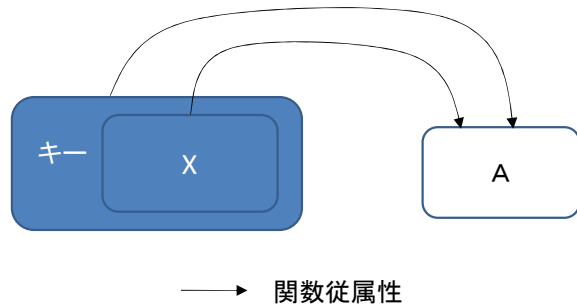
3

完全関数従属性

- 関数従属性 $X \rightarrow Y$ で, X の任意の真部分集合 X' ($X' \subset X$) について $X' \rightarrow Y$ は成立しないとき, Y は X に完全関数従属 (fully functionally dependent) しているという
- 第2正規形における二つ目の条件は重要! 例えば, すべての候補キーがシングル属性の場合は, 二つ目の条件は必ず成り立つ

4

部分的関数従属性



5

更新時異状 (update anomaly, 更新不整合)

注文

顧客名	商品名	数量	単価	金額
A商店	テレビ	3	198,000	594,000
Bマート	テレビ	10	198,000	1,980,000
Bマート	洗濯機	5	59,800	299,000
C社	餅つき機	1	29,800	29,800

- タップル挿入時異状
 - (-, 電子レンジ, -, 74,800, -)を挿入 ← キー制約から, 無理!
- タップル削除時異状
 - (C社, 餅つき機, 1, 29,800, 29,800) ← 重要なデータの喪失!
- タップル修正時異状
 - テレビの単価を198,000 から 148,000に変更 ← 修正大変!
 - C社からの注文を, 餅つき機から洗濯機に変更 ← 重要なデータの喪失!

6

第2正規形でないリレーション

注文

顧客名	商品名	数量	単価	金額
A商店	テレビ	3	198,000	594,000
Bマート	テレビ	10	198,000	1,980,000
Bマート	洗濯機	5	59,800	299,000
C社	餅つき機	1	29,800	29,800

$F = \{f_1: \{\text{顧客名, 商品名}\} \rightarrow \text{数量},$
 $f_2: \text{商品名} \rightarrow \text{単価}, f_3: \{\text{商品名, 数量}\} \rightarrow \text{金額},$
 $f_4: \{\text{数量, 単価}\} \rightarrow \text{金額}, f_5: \{\text{数量, 金額}\} \rightarrow \text{単価},$
 $f_6: \{\text{単価, 金額}\} \rightarrow \text{数量}\}$

主キー(候補キー)は{顧客名, 商品名}
 非キー属性は, 数量, 単価, 金額
 単価は, f_2 より, 商品名(← 候補キー)に関数従属
 \therefore 第2正規形ではない

7

第2正規形でないリレーション

注文

顧客名	商品名	数量	単価	金額
A商店	テレビ	3	198,000	594,000
Bマート	テレビ	10	198,000	1,980,000
Bマート	洗濯機	5	59,800	299,000
C社	餅つき機	1	29,800	29,800

$F = \{f_1: \{\text{顧客名, 商品名}\} \rightarrow \text{数量},$
 $f_2: \text{商品名} \rightarrow \text{単価}, f_3: \{\text{商品名, 数量}\} \rightarrow \text{金額},$
 $f_4: \{\text{数量, 単価}\} \rightarrow \text{金額}, f_5: \{\text{数量, 金額}\} \rightarrow \text{単価},$
 $f_6: \{\text{単価, 金額}\} \rightarrow \text{数量}\}$

f_2 を使って,
 注文[顧客名, 商品名, 数量, 金額]と注文[商品名, 単価]
 に情報無損失分解すると, 第2正規形となる.

8

更新時異状は解消される！

- タプル挿入時異状
 - (, 電子レンジ, -, 74,800, -)を挿入←キー制約から、無理！
- タプル削除時異状
 - (C社, 餅つき機, 1, 29,800, 29,800) ←重要なデータの喪失！
- タプル修正時異状
 - テレビの単価を198,000 から 148,000 に変更←修正大変！
 - C社からの注文を, 餅つき機から洗濯機に変更←重要なデータの喪失！

注文(=注文[顧客名, 商品名, 数量, 金額])

顧客名	商品名	数量	金額
A商店	テレビ	3	594,000
Bマート	テレビ	10	1,980,000
Bマート	洗濯機	5	299,000
C社	餅つき機	1	29,800

商品(=注文[商品名, 単価])

商品名	単価
テレビ	198,000
洗濯機	59,800
餅つき機	29,800

注文＝注文[顧客名, 商品名, 数量, 金額] * 注文[商品名, 単価]

9

第2正規形だが第3正規形でない

社員				
社員番号	社員名	給与	所属	勤務地
0650	山田太郎	50	K55	神奈川
1508	鈴木花子	40	K41	東京
0231	田中桃子	60	K41	東京
2034	佐藤一郎	40	K55	神奈川
2100	高橋次郎	40	K58	静岡

属性名のアンダーラインは主キーを表す

社員番号→社員名
社員番号→給与
社員番号→所属
所属→勤務地

主キーは社員番号

- タプル挿入時異状
 - (, -, -, K45, 千葉) ←無理！
- タプル削除時異状
 - (2100, 高橋次郎, 40, K58, 静岡) ←重要なデータの喪失！
- タプル修正時異状
 - K41の部門の所在地を東京から千葉へ←修正大変！
 - 高橋次郎さんの部門をK58からK55へ←重要なデータの喪失！

これらの異状はなぜ？

- リレーション社員では、非キー属性の勤務地が主キーの社員番号に直接に関数従属しておらず、社員番号に**推移的に関数従属**しているから
1. 社員番号→所属
 2. 所属→勤務地
 3. 社員番号→勤務地

11

情報無損失分解

社員				
社員番号	社員名	給与	所属	勤務地
0650	山田太郎	50	K55	神奈川
1508	鈴木花子	40	K41	東京
0231	田中桃子	60	K41	東京
2034	佐藤一郎	40	K55	神奈川
2100	高橋次郎	40	K58	静岡

属性名のアンダーラインは主キーを表す

社員番号→社員名
社員番号→給与
社員番号→所属
所属→勤務地
(社員番号→勤務地)

(a) リレーション社員 — 第2正規形ではあるが第3正規形ではない例—

社員[社員番号, 社員名, 給与, 所属]				社員[所属, 勤務地]	
社員番号	社員名	給与	所属	所属	勤務地
0650	山田太郎	50	K55	K55	神奈川
1508	鈴木花子	40	K41	K41	東京
0231	田中桃子	60	K41	K41	東京
2034	佐藤一郎	40	K55	K55	神奈川
2100	高橋次郎	40	K58	K58	静岡

(b) 関数従属性 所属 → 勤務地 によるリレーション社員の情報無損失分解

12

情報無損失分解

社員[社員番号, 社員名, 給与, 所属]				社員[所属, 勤務地]	
社員番号	社員名	給与	所属	所属	勤務地
0650	山田太郎	50	K55	K55	神奈川
1508	鈴木花子	40	K41	K41	東京
0231	田中桃子	60	K41	K58	静岡
2034	佐藤一郎	40	K55		
2100	高橋次郎	40	K58		

- タプル挿入時異状
 - (-, -, -, K45, 千葉) ← 無理！
- タプル削除時異状
 - (2100, 高橋次郎, 40, K58, 静岡) ← 重要なデータの喪失！
- タプル修正時異状
 - K41の部門の所在地を東京から千葉へ ← 修正大変！
 - 高橋次郎さんの部門をK58からK55へ ← 重要なデータの喪失！

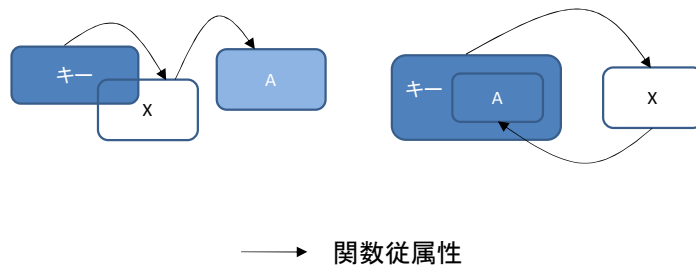
第3正規形

リレーションスキーマRが第3正規形であるとは次の二つの条件を満たすときをいう

1. Rは第2正規形である
2. Rの全ての非キー属性はRのいかなる候補キーにも推移的に関数従属しない

14

推移的関数従属性



15

第3正規形

- 推移的関数従属性が多段に及ぶ場合, たとえば, $R(\underline{A}, B, C, D)$ で $A \rightarrow B, B \rightarrow C, C \rightarrow D$ とあった場合, $R[A, B, C]$ と $R[C, D]$ の後に, $R[A, B], R[B, C]$ の分解が必要
- データベース設計では, 第3正規形にまで正規化するのが普通

16

第2正規形と第3正規形

- 両方とも、関数従属性を見つけて分解することが重要
- 違いは
 - 第2正規形: 候補キーの真部分集合に注目して関数従属性を見つけて、そこで分解
 - 第3正規形: キー以外の部分に注目して関数従属性を見つけて、そこで分解

17

3NFだがBCNFでない



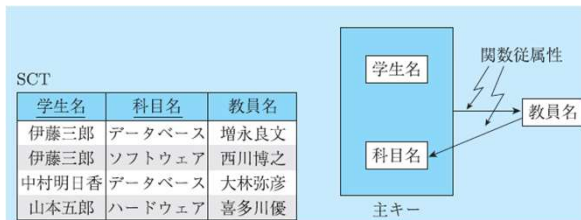
- 1人の学生が同一科目をそれを教授している複数の教員から受けることはない
- ある一つの科目を教えている教員は複数いることはある
- その中の誰か1人の先生を選んで学生は履修する
- 1教員が複数の科目を教えることはできない

候補キー:
{学生名, 科目名}
{学生名, 教員名}

非キー属性がないため、これは第3正規形を満たす

18

3NFだがBCNFでない時の更新時異状



- タップル挿入時異状
 - (., コンピュータグラフィックス, 佐藤祐子) ← 無理!
- タップル削除時異状
 - (伊藤三郎, ソフトウェア, 西川博之) ← 重要なデータの喪失!
- タップル修正時異状
 - ソフトウェアの担当を西川博之から青木康に変更 ← 修正大変!
 - (伊藤三郎, ソフトウェア, 西川博之) を (伊藤三郎, ハードウェア, 喜多川優) に修正 ← 重要なデータの喪失!

19

ボイスコッド正規形

リレーションスキーマRがボイス-コッド正規形であるとは次の条件が成立するときをいう:

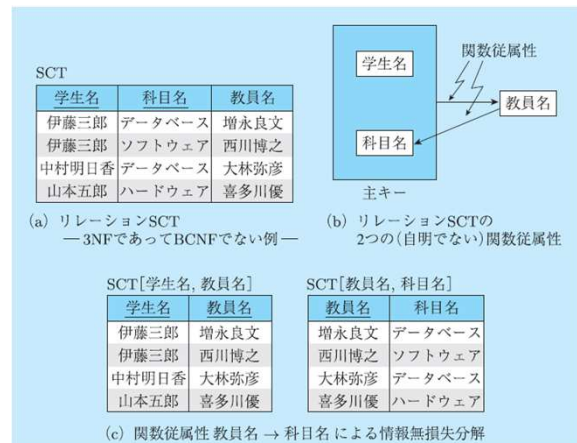
$X \rightarrow Y$ をRの関数従属性とするとき

1. $X \rightarrow Y$ は自明な関数従属性であるか、または
2. XはRのスーパーキーである

注: 自明な関数従属性とは、 $X \rightarrow Y$ を関数従属性とするとき、 $Y = \emptyset$ か $Y \subseteq X$ のときをいう

20

ボイスコード正規形



21

第4正規形

フライト

フライト番号	クルー名	乗客名
55	P	A
55	S	A
55	P	B
55	S	B
55	P	C
55	S	C
505	P'	A'
505	S'	A'

• タップル挿入時異状

- 61便のクルーP'のデータ挿入無理
- 55便のクルーS'のデータ挿入大変

• タップル削除時異状

- 505便のA'の予約のキャンセル大変
- 505便のクルーのデータがなくなり困る

• タップル修正時異状

- 55便のクルーPがP'に変更大変
- 505便の乗客A'が55便に変更困る

これらの解消には、多値従属性
フライト番号 → クルー名 | 乗客名 を使用して
フライト[フライト番号, クルー名]と
フライト[フライト番号, 乗客名]に分解
すればよい

22

第4正規形

リレーションスキーマRが第4正規形であるとは次の条件を満たしているときをいう:

$X \twoheadrightarrow Y$ をRの多値従属性とするととき

1. $X \twoheadrightarrow Y$ は自明な多値従属性であるか、または
2. XはRのスーパーキーである

注: 自明な多値従属性とは、 $X \twoheadrightarrow Y | Z$ を多値従属性としたとき、 $Z = \emptyset$ か $Y \subseteq X$ のときをいう

23

第5正規形

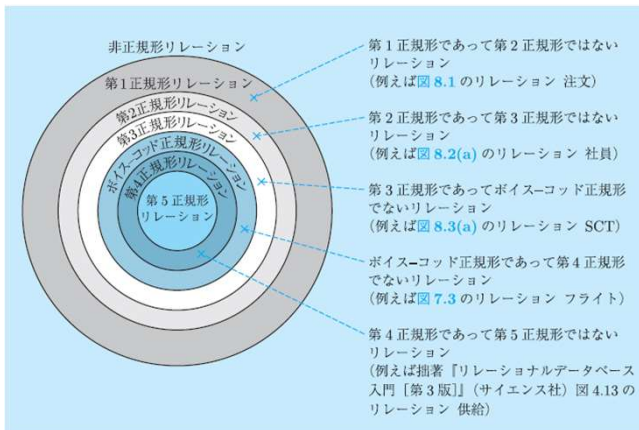
• 結合従属性

- 2分解可能ではないが、3分解可能
- 一般に、n分解可能なものあり
- 自然結合をとると元に戻る従属性
- 2分解の場合、多値従属性

• 自明でない結合従属性をなくすと第5正規形

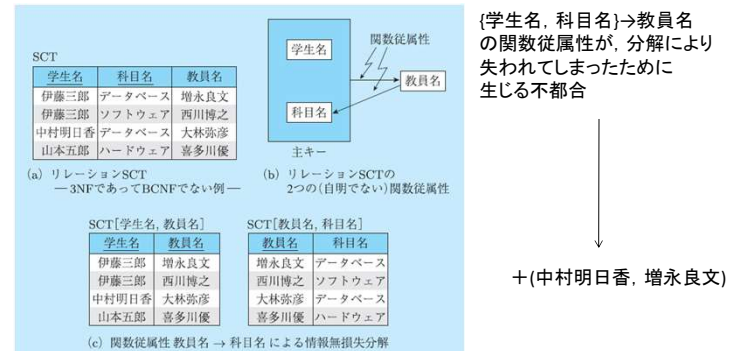
24

正規形の階層関係



25

関数従属性損失分解



26